IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

CONNECTU LLC,

Plaintiff,

v.

MARK ZUCKERBERG, EDUARDO SAVERIN, DUSTIN MOSKOVITZ, ANDREW MCCOLLUM, CHRISTOPHER HUGHES, and FACEBOOK, INC.,

Defendants.

MARK ZUCKERBERG and FACEBOOK, INC.,

Counterclaimants,

v.

CONNECTU LLC,

Counterdefendant,

and

CAMERON WINKLEVOSS, TYLER WINKLEVOSS, and DIVYA NARENDRA,

Additional Counterdefendants.

CIVIL ACTION NO. 1:04-cv-11923 (DPW)

District Judge Douglas P. Woodlock

Magistrate Judge Robert B. Collings

PLAINTIFF CONNECTU LLC'S REPLY TO FACEBOOK DEFENDANTS' RESPONSE TO CONNECTU'S "FORENSIC RECOVERY" ARGUMENTS

I. INTRODUCTION

Facebook Defendants' Opposition, which ignores most of its own admissions and many of ConnectU's arguments presented at the March 3, 2006 hearing and in ConnectU's March 10, 2006 "Summary" thereof (Docket No. 148), confirms ConnectU's position. Specifically, Facebook Defendants admit that they did not do all that

ConnectU would do to find the missing source code, they admit that they neither imaged nor forensically examined any memory devices before November 18, 2005, and they admit (or do not seriously dispute) that the representations made by Facebook Defendants' counsel during the November 18, 2005 hearing were not true. Desperate to avoid producing their memory devices and the images thereof, Facebook Defendants' engaged another expert after the March 3, 2006 hearing to search once again the memory device images they made after the November 18, 2005 hearing, thereby conceding that their initial search was inadequate. Facebook Defendants' inadequate "do-over" is too little too late, as discussed below. ConnectU therefore urges the Court to allow ConnectU's expert to conduct an independent forensic examination of memory Device Nos. 371-01 through 371-10[1] and the images that Facebook Defendants made of them, and to order Facebook Defendants to provide the other discovery requested at the March 3 hearing and in ConnectU's Summary.

II. DISCUSSION

In an effort to file a concise brief, ConnectU will address only the most egregious and important problems with Facebook Defendants' Response and supporting declarations.

A. <u>Facebook Did Not Do Everything ConnectU Would Do</u>

ConnectU wants to find the missing code, specifically: (1) the portion of the Harvard Connection Code allegedly written by Defendant Zuckerberg, (2) the complete

[1] ConnectU understands that there are other devices in the possession, custody, and/or control of the Defendants that have not been images or searched. (*See e.g.* Facebook Defendants' response to Interrogatory No. 25 at p. 8, noting the existence of Penguin servers, which were not searched.). This Motion does not address such devices. ConnectU may request images of such devices in a separate motion.

Facebook code, including the database definitions, from before, at, and after the time of launch (i.e. February 4, 2004), (3) the coursematch code, and (4) the facemash code. Facebook Defendants do not and cannot dispute that each of these source codes existed and that they were stored on at least Mr. Zuckerberg's memory devices at one time. Facebook Defendants have now engaged two experts to search for these codes, only one of which (Mr. Butterworth) submitted a declaration that the codes cannot be found using the search methods employed. Facebook Defendants have now had two shots at thoroughly searching these devices, and neither searcher did everything ConnectU would do. Rather than being forced to accept Facebook Defendants' experts' representations, ConnectU's expert should be permitted to examine these devices (or at least the forensic images allegedly created by Facebook Defendants' experts) and form his own opinion regarding the code and related information that they do or do not contain, and, if the code cannot be found, what happened to it.

ConnectU's March 3 presentation and March 10 Summary gave numerous examples of how Facebook Defendants' first expert did not do what ConnectU's expert would have done. They were only examples, as ConnectU said (March 3, 2006 Hearing Transcript at v. 1 at p. 9-13 (Docket No. 151) (hereinafter "Trans."); Plaintiff's Summary at p. 8-10). Facebook Defendants argue that their new expert, Mr. Butterworth, either did what ConnectU's expert would have done, or that doing so is too expensive. Responding to the latter argument first: It would cost Facebook Defendants virtually nothing to produce the devices and allow ConnectU's expert to image and search them. Thus, their argument that such a search is overly burdensome is without merit. And although Mr. Butterworth chose not to perform certain forensic processes, ConnectU would have performed them and the test set by the Court on November 18, 2005 was the latter, not the former.

More importantly, the examples of the shortcomings of Facebook Defendants' forensic analysis presented by ConnectU at the March 3, 2006 hearing were just that: examples. In addition to such examples, ConnectU's expert also would have, *for example*: (1) searched for evidence of intentional overwriting, (2) determined how much of Device No. 371-01 is unreadable, (3) recovered evidence remaining in the readable portions[2], (4) conducted a content-based -- rather than file-extension-based -- search of slack space, unallocated clusters, uncommitted disk sectors (the largest part of most disk drives), and existing files, which could show that Facebook Defendants' deleted the code, and (5) conducted other analyses suggested by the images of the devices.[3] Perhaps most importantly, Facebook Defendants' experts have merely, but imperfectly, searched for the missing code, but have made no effort to explain why such code cannot be found. And why would they? After all, they work for Facebook Defendants.

For these reasons, Facebook Defendants should be ordered to produce the images they made of Facebook Defendants' memory devices, and the devices

[2] Note that the Butterworth declaration at ¶ 5 confirms that a "rudimentary search" was possible but not conducted. Indeed, Mr. Butterworth's declaration fails to note even basic information, such as the total number of sectors that could not be read or the total number of bytes represented by that data. Moreover, Mr. Butterworth limited his analysis to Mr. Berryhill's image. Thus, any shortcomings of Mr. Berryhill's sloppy work were inherited by Mr. Butterworth, whose analysis was therefore limited by Mr. Berryhill's incomplete image. Facebook Defendants have expressed interest in surreplying to this brief (which ConnectU opposes), so its do-overs can continue *ad nauseum*.

[3] It is impossible to identify each step of the investigation. By their nature, forensic investigations, like any investigation, rely not only on carrying out the basic protocols, but to follow leads as they arise.

themselves, for imaging and evaluation by ConnectU's expert, at Facebook Defendants' expense. Facebook Defendants should also be ordered to provide the other discovery requested at the March 3 hearing and in ConnectU's Summary.[4]

B. Facebook Defendants Withhold Key Evidence

Facebook Defendants have admitted that Device No. 371-01 was the computer Mr. Zuckerberg used at Harvard during the 2003-2004 academic year (Max Kelly Depos. Trans. at p. 66-67).[5] As mentioned above, Facebook Defendants make no effort to explain why the missing code cannot be found or what happened to it. They have also produced no evidence regarding when, how, and where Mr. Zuckerberg used his various devices, relying entirely on unsupported attorney argument, which contradicts Facebook Inc.'s own Rule 30(b)(6) witness, that the 371-01 device "was not given to Facebook's counsel until late October 2005 because Facebook personnel knew that the device had malfunctioned . . . ." (Response at p. 15-16) Facebook Defendants do not even attempt to back-up this attempt to explain away their now-conceded spoliation and suppression of evidence with any proof.[6] Even where documentary

[4] For example, Facebook Defendants argue that none of ConnectU's arguments suggest that their responses to Interrogatory Nos. 25-26 were inaccurate (Response at 15). To the contrary, ConnectU raised the inadequacy of such interrogatory responses at the hearing (Trans. at p. 16) and in the Summary at 16-22.

[5] Facebook Defendants attempt to obscure this admission with double talk, introducing co-Defendant McCollum's hard drive, Device No. 371-05 (Response at 15-16). What Facebook Defendants' explanation of the relationship between Device Nos. 371-01 and 371-05 lacks, besides any clarity, is evidence to support it.

[6] Facebook Defendants' Rule 30(b)(6) witness's (Mr. Max Kelly) testimony and Facebook Defendants' other admissions, presented by ConnectU at the March 3 hearing and in ConnectU's Summary, entail suppression and spoliation of evidence. Mr. Kelly attempted to change his unequivocal testimony in the errata to his deposition (attached hereto as Ex. A). Now that they have been caught red-handed in the

(continued on next page)

evidence exists, Facebook Defendants admit withholding it. For example, the Response (p. 6) and the Butterworth Declaration (¶ 6.d) both state that Facebook Defendants' expert collected metadata for selected files, the results of which were "set out in spreadsheets." No such spreadsheets were produced. If ConnectU's expert had forensically analyzed Facebook Defendants' images and devices, ConnectU would have had this information months ago.

Facebook Defendants also continue to withhold the crucial database definitions ("DBDs"), which are an intrinsic part of the database used by the Facebook website. Facebook Defendants cannot deny they exist (either now or at the time of launch). If there is no separate file of the DBDs, they exist in the database itself, and can be easily exported or printed. Without them, the Facebook code is incomplete and useless for a copyright infringement code comparison.

Amazingly, Facebook Defendants continue to play costly word games regarding its DBDs (Response at p. 8, penultimate paragraph), then in the next paragraph concede that they know what the term means ("Guidance has confirmed that what it

(continued from previous page)
suppression and spoliation of evidence, based on their own inescapable admissions, Facebook Defendants want to change sworn 30(b)(6) testimony, interrogatory answers attested to by Mr. Max Kelly (see Response at n. 6) (although they have not yet done so), and present a confusing if not nonsensical explanation without any evidence to support it. But such changes and such arguments, which contain several new admissions (Response at p. 15-16), confirm such suppression and spoliation (e.g, Facebook Defendants say "the second owner of the laptop may have re-formatted the hard drive then used in Device 371-05 when he took possession of the laptop." (Opp. at 16) Co-Defendant Andrew McCollum owned Device No. 371-05 (see Response to Interrogatory No. 25 at p. 9 -- "Andrew's Laptop #1"). If he took possession after facebook.com launched, Mr. Zuckerberg failed to preserve evidence and Mr. McCollum deleted crucial evidence when he reformatted. No matter how Facebook Defendants try to explain it, they suppressed and destroyed evidence. However, ConnectU will explain these issues in greater detail in its motion for sanctions.

understands are database definition files act or, more appropriately, can act, as a reference file for how a database should create tables which record information.”). They also argue that ConnectU has not explained the relevance of the DBDs (Response at p. 8) when it has done so over and over and over again.

Facebook Defendants’ counsel admitted at the March 3, 2006 hearing that the DBDs would be on facebook.com’s production servers. (Trans. at p. 24) Facebook Defendants have not said whether they looked on such servers. However, the DBDs from prior to launch would have been on Mr. Zuckerberg’s personal computer (i.e., 371-01, the device Facebook Defendants have unequivocally admitted was the computer he used at Harvard during 2003-2004 (Max Kelly Depos. Trans. at p. 66-67). The DBDs from launch would have been on Mr. Zuckerberg’s computer and the launch server, at least, and probably also on Defendant Moskovitz’s device(s) (Facebook Defendants have produced no data from its launch servers). Again, Facebook Defendants ask ConnectU to identify facebook.com’s DBDs, but only they know how THEY named them, and therefore how to find them. ConnectU’s burden is simply to ask for the DBDs for facebook.com from prior to launch, launch, and up through December 31, 2004.

C. **ConnectU’s Requests are Legitimate E-Discovery Procedures**

Mr. Butterworth declares at ¶ 11 that searching slack space is normally limited to criminal cases, and is “burdensome, expensive and, virtually always, fruitless in civil matters.” This cursory conclusion is unsupportable. File slack is the unused portion of a file resulting from the fact that the operating system allocates space for files in fixed increments, which usually exceed the actual amount required. File slack will generally contain fragments of files previously written to that disk space. In this case, the presence of source code in file slack would be probative of the issue of whether relevant source code was overwritten or otherwise deleted from the device. With respect to Device No. 371-01, even a fragment would prove what everyone knows: the code was stored on the computer at one time. Therefore, a search of file slack for source code should be included in the forensic examiner's analysis and is something ConnectU would have done.

Mr. Butterworth offers a scenario relevant to a criminal case, in which a victim's bank account number is found in the slack space of an unrelated file, but says he doesn't see that technique as useful in a civil matter. But the computer might assign to a memory cluster that once held a missing source code file a smaller file of some other type. That portion of the memory cluster that was not overwritten by the smaller file would continue to hold remaining program code, and a content-based file slack search could reveal it. Information in file slack often survives longer than information in unallocated clusters, because file slack will remain unchanged for as long as the successor file holds that space, whereas unallocated clusters (discussed below) are always subject to reassignment as space for a new file. For example, if the court were to delete a document of 100KB from its computer, such 100KB of memory becomes available for another use. If the court then creates a new document of 75KB and the computer assigns it to such memory space, 25KB (25%) of the deleted document could be recovered for as long as the court keeps the 75KB document. Accordingly, file slack searching can be very important in civil cases.

Defendants' counsel attributes to Mr. Butterworth (although Mr. Butterworth made no such statement) a conclusion that "analysis of file slack space . . . must be done entirely manually." (Response at 9) This is not true. Used correctly, EnCase forensic software will simultaneously search active files, unallocated clusters, and file slack for any number of designated search terms. The manual part of the process is the examiner's review of discovered hits to determine which are of value. Further, this burden of manual review--if indeed it is a burden at all--would fall to ConnectU if Facebook Defendants produce the devices and/or images for examination.

Mr. Butterworth also declares at ¶ 11 that searching unallocated clusters is normally limited to criminal cases, not civil discovery. This opinion is similarly unsupportable. Unallocated clusters consist of disk space freed up by the deletion of files or space which has never been used. To omit the searching of unallocated clusters would mean not searching for deleted files. As indicated above, where, as here, a motive to delete data exists, a normal forensic step would be to search for that data in unallocated clusters.

Further, Mr. Butterworth's rationale for not searching unallocated clusters because only file fragments would be found there is misleading because file fragments are typically several clusters in length. A single cluster is at least 4,096 bytes, and often more. In terms of PHP code, a cluster would generally contain at least three pages of code (possibly much more), so finding one or more clusters would be very effective in looking for portions of the deleted Facebook code. Using the incomplete October 2004 version of the Facebook code (which Facebook Defendants produced) as an example, the majority of the PHP files in that collection would fit into an 8,192-byte fragment. In other words, most of the PHP code files sought by Plaintiff would not be fragmented at all.

The International Association of Computer Investigative Specialists (IACIS) has published a guideline *Forensic Examination of Computers and Digital and Electronic Media*. In the section "Guide for Forensic Examinations," it recommends that unused and unallocated space on each volume should be examined for previously deleted data, deleted folders, slack space data, and intentionally placed data. *See* http://www.iacis.info/iacisv2/pages/forensicprocedures.php. Guidance Software covers the same principle in its training programs, and the EnCase Forensic software User Manual provides instructions for such searches.

D. <u>Defendants' Production Precludes Reliable Forensic Examination</u>

Information found on a computer hard drive can be produced in a myriad of ways. Three are at issue here. ConnectU's first preference is to obtain the memory devices themselves to allow ConnectU's expert to create and analyze forensic images of such devices. This approach takes ConnectU directly to the source of the data. Alternatively, but not as reliable from a forensic viewpoint, ConnectU wishes to obtain the hard drive images created by Facebook Defendants. Because Facebook Defendants' first expert's (Mr. Berryhill) work was so sloppy, as Facebook Defendants concede, ConnectU can't be certain that the images were properly made. Even more forensically unreliable are the CDs produced by Facebook Defendants, onto which Mr. Berryhill selectively copied files found on the images he made. As detailed in ConnectU's Summary, these CD copies contain virtually no useful information and in fact contain misinformation or disinformation because the process of copying files from the image to another medium (whether a CD or DVD) changed the information; in particular, all directories are created with new dates, and most date-timestamp information is lost for files. Mr. Butterworth's declaration confirms that the information provided on the CDs lacks needed file attributes. His examples at ¶ 12 show that the EnCase image includes numerous file attributes, including File Creation Date, Last Written Date, Entry Modified Date, and Last Accessed Date. Facebook Defendants admit that they have this information (Response at p. 11, first full paragraph).

When files are copied onto a CD, the Last Modified Date of each file copied is the only date carried over to the CD. For example, at the hearing, ConnectU showed two facebook.com code files last modified December 22 and 23, 2003 (see Summary, Ex.9). As Mr. Butterworth confirmed, only Device No. 371-09 or the image made from it will show the other attributes of these files (and Facebook Defendants are withholding this information because they are withholding the devices and images).[7] See Butterworth Decl. at ¶ 12. In this regard, Facebook Defendants misstate ConnectU's argument (Response at p. 11) because they talk of COPYING files to folders. Copying existing files will show the date of copying. But ConnectU seeks data on file and directory CREATION, which the image will show.

When files were copied from the images to the CDs, other relevant information was also lost by the copying process, such as (a) deleted files in unallocated sectors; (b) partial files in file slack; and (c) temporary files or logs created by applications or by the operating system that may hold important information.

[7] Facebook Defendants seems to distinguish metadata and file attributes (Response at p. 11, first full paragraph), then ask ConnectU to identify the metadata it wants. To the extent this is not just another confusion technique, ConnectU's response is that it seeks the devices and the images themselves so that its expert can analyze the metadata and file attributes to form relevant opinions, such as what happened to the missing code, when Mr. Zuckerberg first started working on facebook.com, and the extent to which evidence has been spoiled and suppressed.

Because such information is unavailable to ConnectU and Facebook Defendants have no incentive to look for it or analyze it, ConnectU's expert's ability to form opinions about when a device was in use or out of use, when a device was allegedly missing or malfunctioning, when a device was reformatted, whether Facebook source code and database files were deleted, whether software was used to scrub a hard drive, and whether relevant data was overlooked by Facebook Defendants' narrow search methods, is severely hampered.

ConnectU should be entitled to have its expert examine the same evidence Facebook Defendants' two experts have now examined (i.e., the devices and the images) to form his own opinions (a) on the substantive issues, rather than having to take Facebook Defendants' word for it, and (b) as to whether Facebook Defendants' experts are correct. If Facebook Defendants are ordered to produce the devices and images, ConnectU's expert can see such data, rather than forcing ConnectU to guess the files for which it needs such information, specifying them to Facebook Defendants, and trusting them to provide the correct and complete data.

For these reasons, Facebook Defendants should be ordered to produce either the devices themselves or at least the images created by the forensic experts.[8]

E. <u>ConnectU Produced the Entire Harvard Connection Code</u>

Facebook Defendants' argue in their Opposition that ConnectU has in some way been remiss in its production. (Response at p. 17) Initially, ConnectU notes that there

[8] Regarding Device No. 371-01, which is the device owned and used by Defendant Zuckerberg at the time the dispute arose, and which allegedly has some unknown sort of data errors, Defendants should be ordered to produce the device for analysis by ConnectU's expert because the cause, timing, correctability, and effect of such alleged corruption is also in issue.

is no motion pending regarding its production and to the extent that Facebook Defendants believe that a document request has not been fully responded to, it is under an obligation to initiate the meet and confer process, which it has not done. In any event, Facebook Defendants admit that ConnectU produced the Harvard Connection code and its DBDs when they state that they received SQL tables. Id. As the Facebook Defendants readily admit, the Harvard Connection Code DBDs are located on CD C011338: D:\connectu-archives\20040101-client_supplied_site\mysqldump.sql, which ConnectU produced almost eight months ago.

F. ***Subversion* File Proves that Facebook's Production is Inadequate**

Facebook Defendants are so eager to avoid producing the devices and images that they inadvertently proved that their search for and production of the missing code is woefully inadequate. Specifically, in attempting to prove that Facebook Defendants did not spoil evidence by including deletion software on the 371-02 (Maverick Server) device, they admitted that the device contains Subversion-brand software.[9] The Subversion program or its database was never produced to ConnectU (or even mentioned), and until Facebook Defendants noted it in their brief, ConnectU had no idea it existed.[10]

[9] "Subversion" is a software product that is used for code version and revision control. See http://subversion.tigris.org.

[10] ConnectU found a file entitled backup.pl on Device No. 371-02. The purpose of this file is to create a backup archive (a consolidated copy) of the website and databases, then copy that archive to a separate server; it then deletes the prior cycle's backup archive. Facebook Defendants' new admission that Device No. 371-02 contains Subversion software is far more nefarious. This is another good example of why Facebook Defendants cannot expect ConnectU to "tell them what we want" (see Response at pp. 1, 10, 11).

As Facebook Defendants note, the Subversion software allows software developers to maintain all versions of their code. (Response at p. 13) It does this by creating a database, which acts as a repository of all versions of the code and maintains a record of exactly what has changed from version to version, allowing a user, *inter alia*, to keep track of prior versions with dates and authors for every change, compare prior versions to the current version, and recall any chosen version for review or reuse.

During the Rule 30(b)(6) deposition of Facebook, Inc. regarding the forensic discovery issues, Mr. Kelly testified that the Maverick server (Device Nos. 371-02 and 371-03) was purchased sometime in 2004. (Max Kelly Depos. Transcript at p. 82-83) He further testified that this server was the primary source code repository until June 2005, after which the code was moved to a different server. Id. Because the Subversion software was found on Device No. 371-02, Defendant Facebook must have a Subversion database containing all the code housed on the 371-02 device from the time it was placed in service in 2004 through June 2005. When a code file is moved into a Subversion database, its file name extension is no longer visible, as the Subversion database holds all data in an independent native format. Facebook Defendants did not search or produce the Subversion database. Yet the very purpose of Subversion is to hold and manage source code, the explicit object of Plaintiffs' production request. Moreover, Mr. Kelly testified that the code was moved to a different server in June 2005, implying that the new server was populated with code written in 2004. Facebook has not identified or searched this new server. Although numerous servers are mentioned in Defendants' report of the search for potential code sources, none except Maverick was delivered to their experts for imaging.

Because Facebook Defendants attempted to conceal the existence of a Subversion database by not producing the file in their meager CD production, this Court should order Facebook Defendants to produce the devices for imaging by ConnectU's expert.

G. Defendant Facebook's California State Action is Irrelevant

For reasons beyond ConnectU's understanding, Facebook Defendants' counsel believes that the instant action is somehow related to Facebook, Inc.'s lawsuit in California. Other than the parties, the suits could hardly be more different. ConnectU's counsel of record are different in the two cases. Thus, ConnectU fails to see the relevance of the deposition schedule and topics in the California action. Why would ConnectU's counsel in the California action want to waste deposition time asking about forensic discovery in this case? In any event, we understand that the deposition of Mr. Zuckerberg noticed in the California action has been postponed at Mr. Zuckerberg's request.[11]

III. CONCLUSION

For the reasons detailed in ConnectU's March 3, 2006 report and summary thereof, and this Reply, ConnectU urges the Court to order Facebook Defendants to produce Device Nos. 371-01 through 371-10, and the images they made of them, and to provide the other discovery requested at the March 3 hearing and in ConnectU's Summary.

[11] ConnectU engaged Finnegan Henderson's California office to handle the California litigation. None of the attorneys in the instant case are of record in that case.

DATED: April 10, 2006

/s/ John F. Hornick
Lawrence R. Robins (BBO# 632610)
Jonathan M. Gelchinsky (BBO# 656282)
FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.
55 Cambridge Parkway
Cambridge, MA 02142
Telephone: (617) 452-1600
Facsimile: (617) 452-1666
larry.robins@finnegan.com
jon.gelchinsky@finnegan.com

John F. Hornick (*pro hac vice*)
Margaret A. Esquenet (*pro hac vice*)
Troy E. Grabow (*pro hac vice*)
FINNEGAN, HENDERSON, FARABOW,
GARRETT & DUNNER, L.L.P.
901 New York Avenue N.W.
Washington, DC 20001
Telephone: (202) 408-4000
Facsimile: (202) 408-4400

Attorneys for Plaintiff and Counterclaim Defendants

CERTIFICATE OF SERVICE

I hereby certify that this document filed through the ECF system will be sent electronically to the registered participants as identified on the Notice of Electronic Filing (NEF) and paper copies will be sent to those indicated as non registered participants on April 10, 2006.